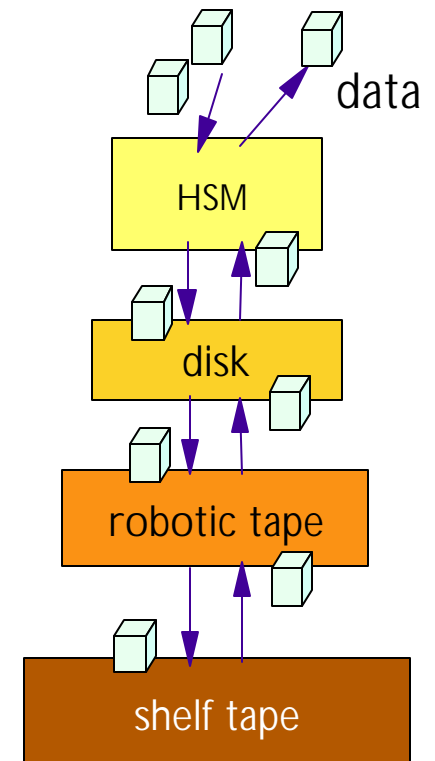


# What is HPSS?

- ❁ High Performance Storage System (HPSS) is software that provides hierarchical storage management (HSM) and archive functions for high-performance computing environments
- ❁ Manages hierarchy of disk arrays, disks, tapes
- ❁ Automatic data migration between levels

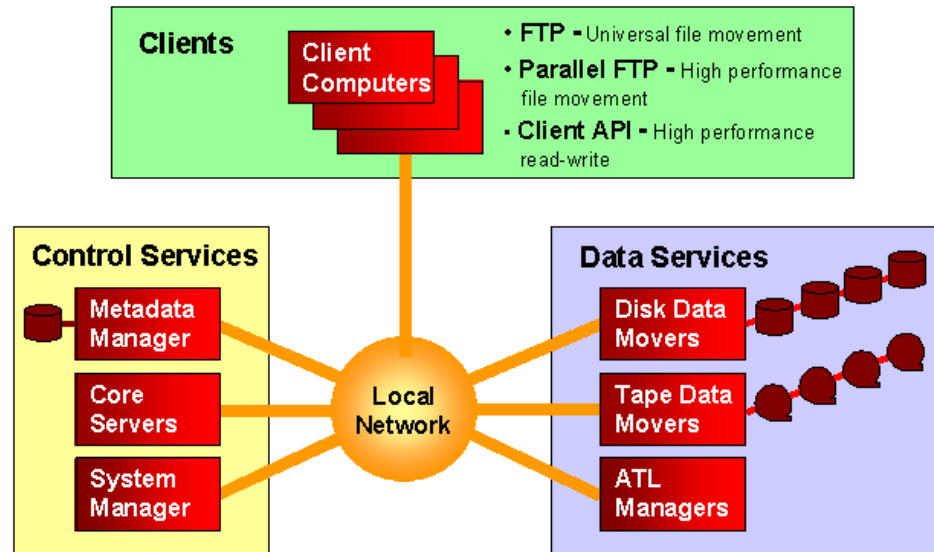


# What is HPSS?

- The product of a collaboration between IBM and the Department of Energy. Developers/Owners are:
  - IBM Global Services - Federal
  - Lawrence Livermore National Laboratory
  - Los Alamos National Laboratory
  - Sandia National Laboratory
  - Oak Ridge National Laboratory
  - Lawrence Berkeley National Laboratory
  - NASA Langley Research Center
- An IBM commercial offering
- Open system with source code available to licensees

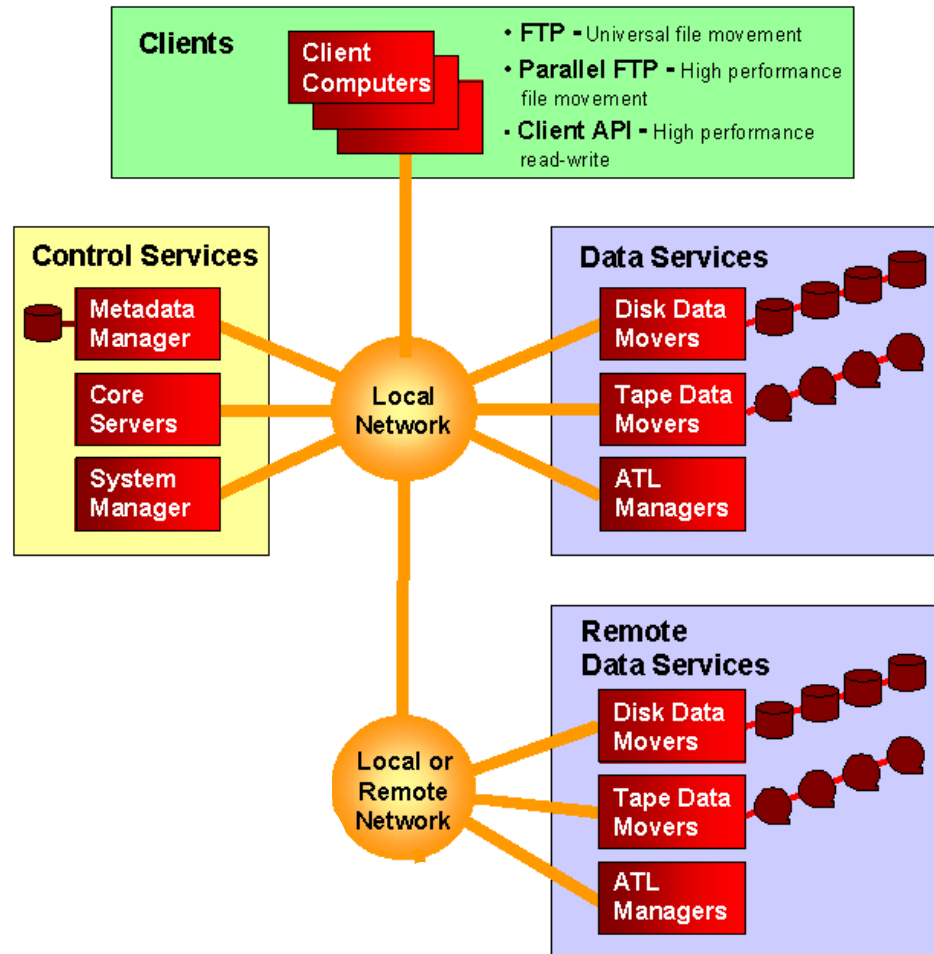
# What is HPSS?

- Network-centered design
- Very fast for large files
- Scalable architecture



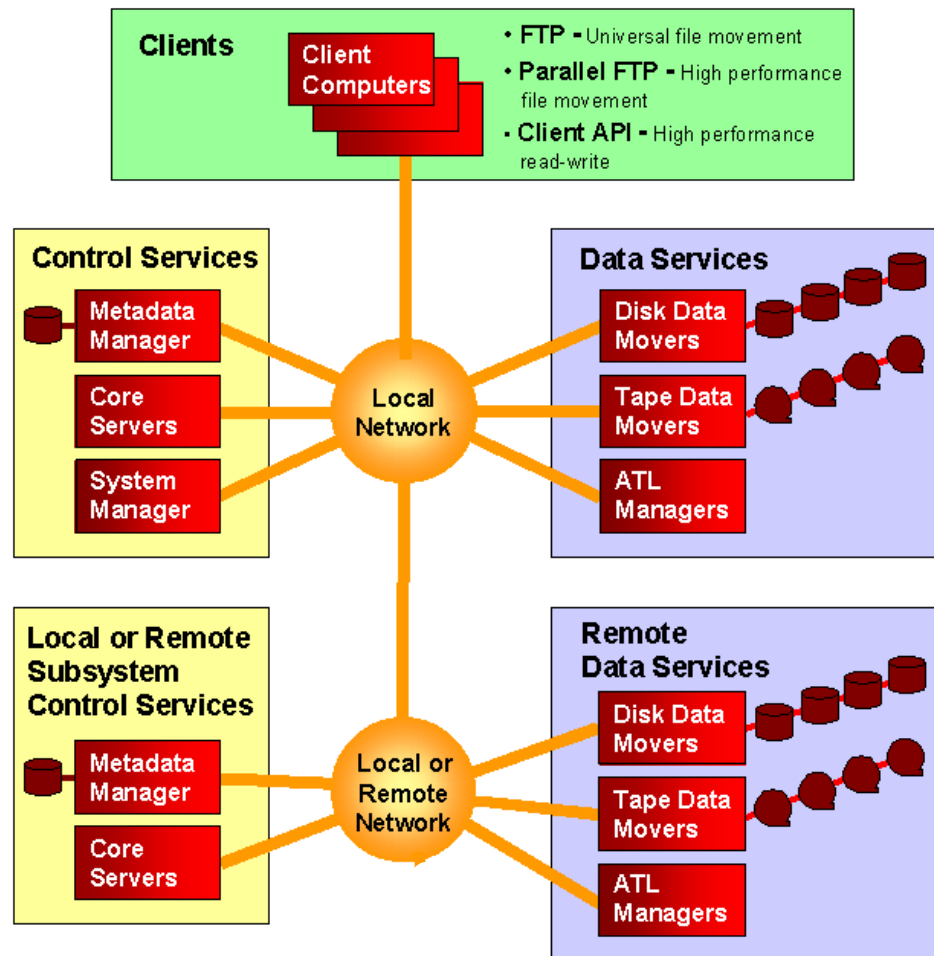
# What is HPSS?

- Network-centered design
- Very fast for large files
- Scaleable architecture
- Remote movers



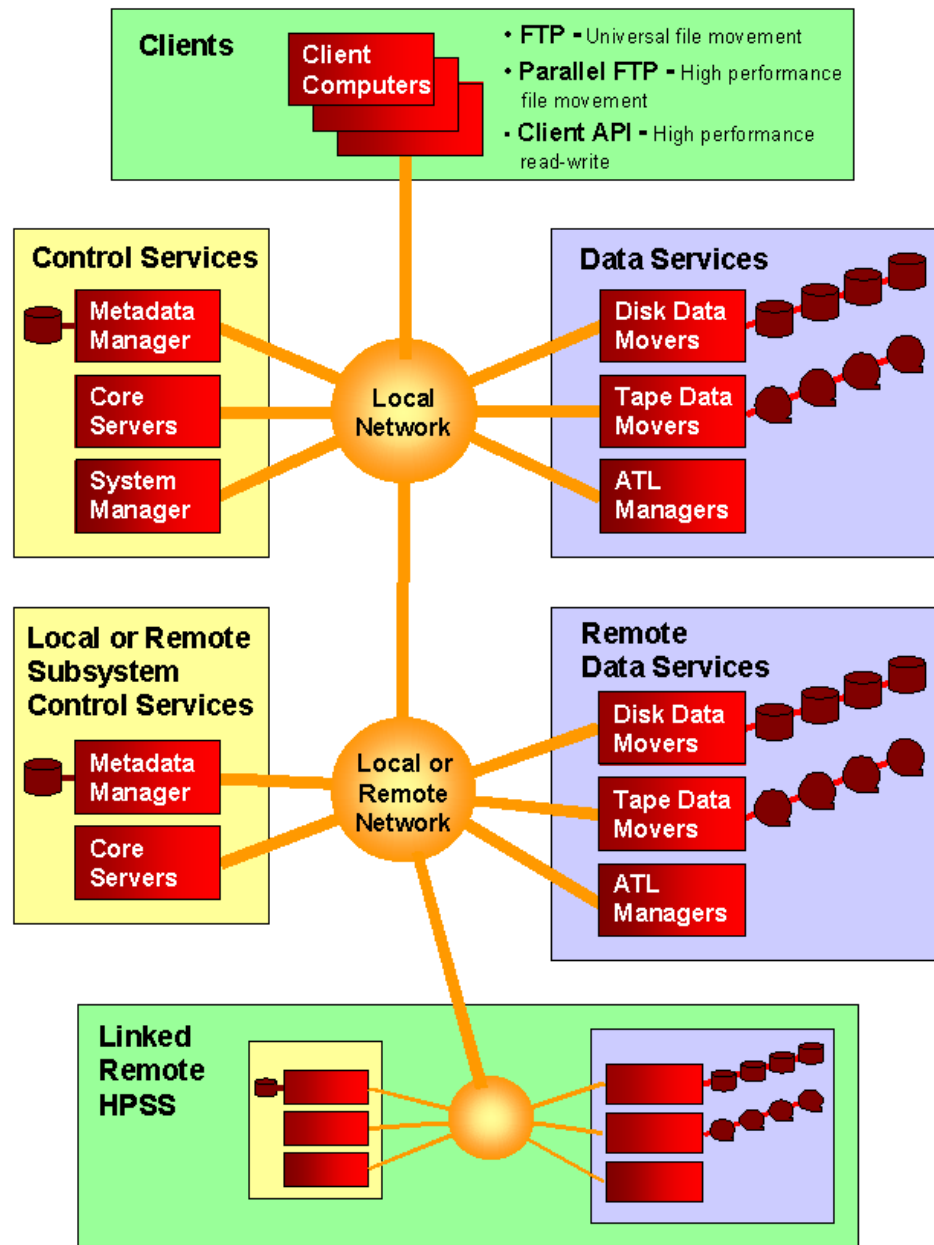
# What is HPSS?

- Network-centered design
- Very fast for large files
- Scaleable architecture
- Remote movers
- Subsystems

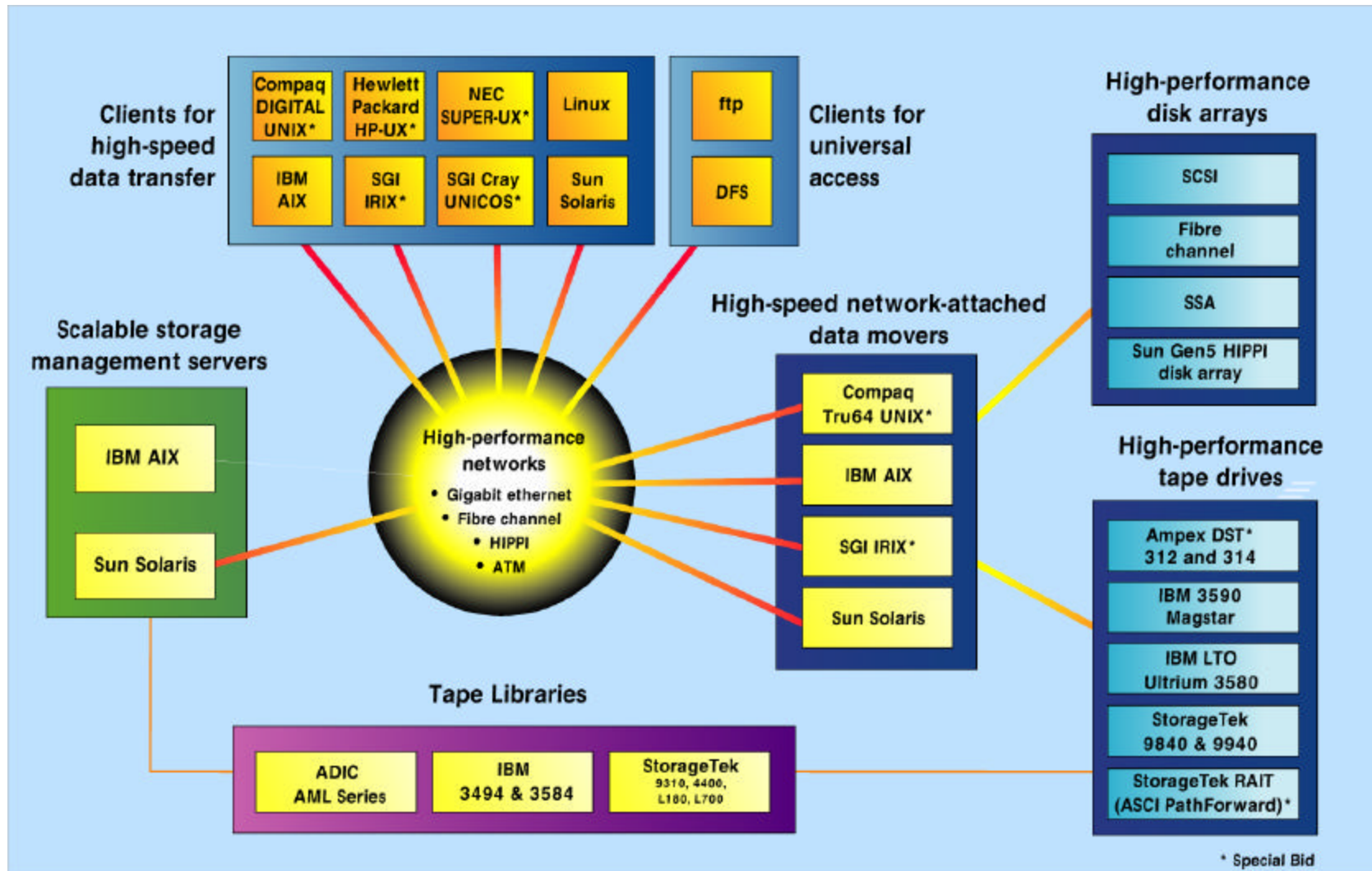


# What is HPSS?

- Network-centered design
- Very fast for large files
- Scaleable architecture
- Remote movers
- Subsystems
- Linked remote HPSS



# HPSS: A Multi-Vendor Mass Storage Solution



# The World of HPSS

## HPSS Deployment Partners





# Enhanced Scalability

## ❁ Why worry with Scalability?

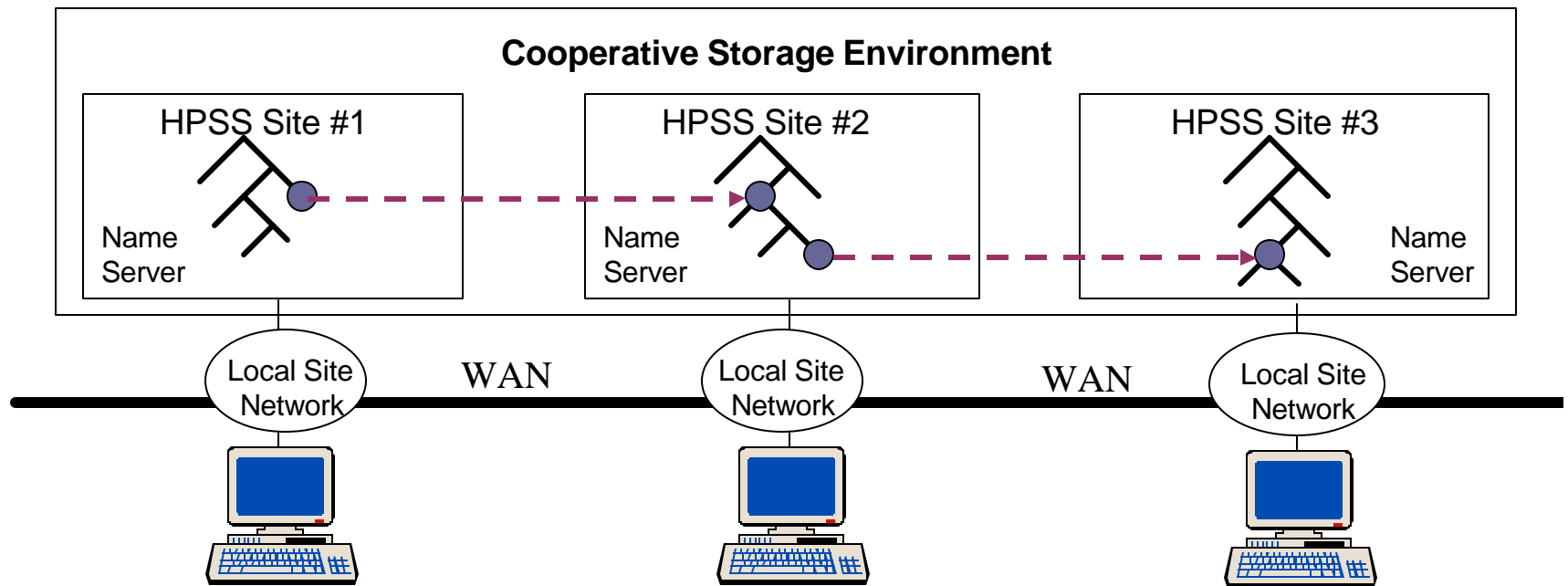
- Petabyte systems are coming
- The DOE's Accelerated Strategic Computing Initiative (ASCI) has aggressive scalability requirements
- Our goal is to be the premier high-end storage system, **period**.

## ❁ Scalability and HPSS

- Huge file systems
- Parallel I/O
- Distributed movers
- Storage subsystems
- Collaborative HPSSes

# Subsystems and Linked Remote HPSS

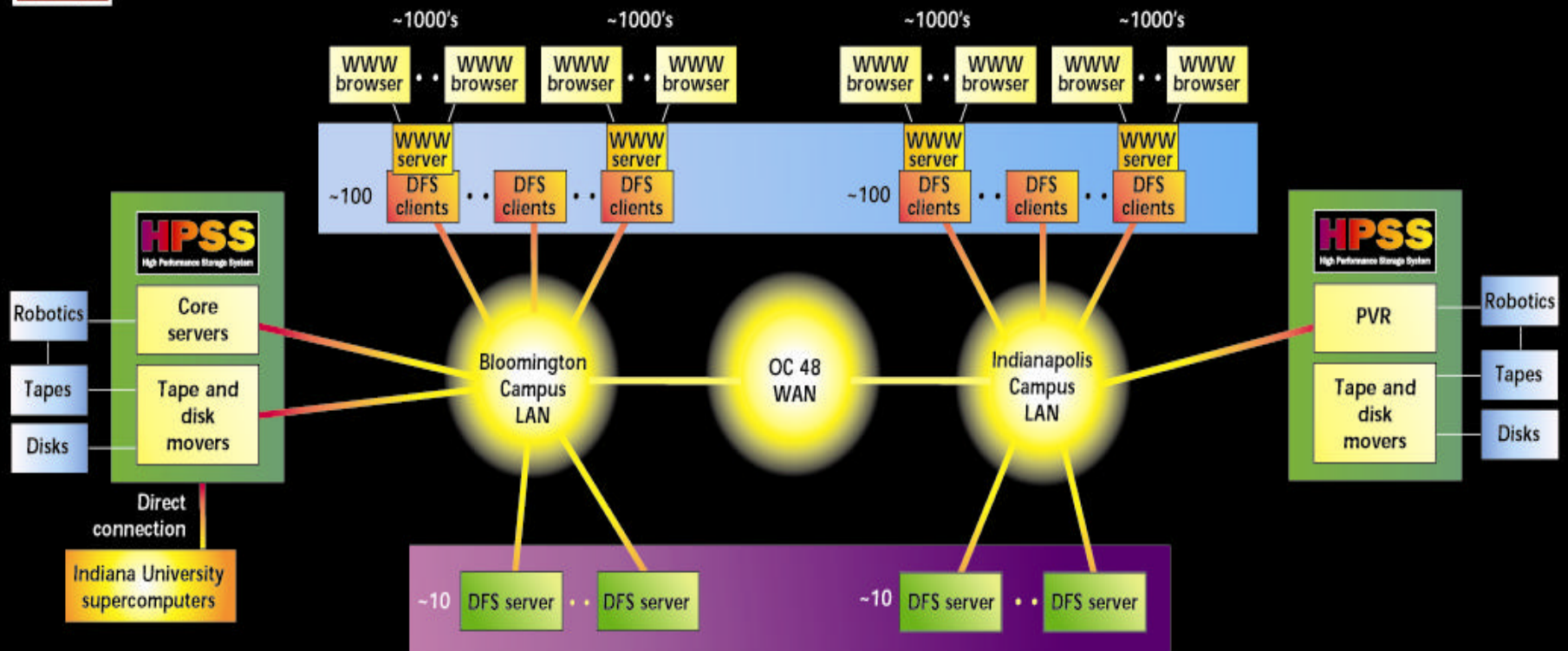
- Federated Name Space Diagram





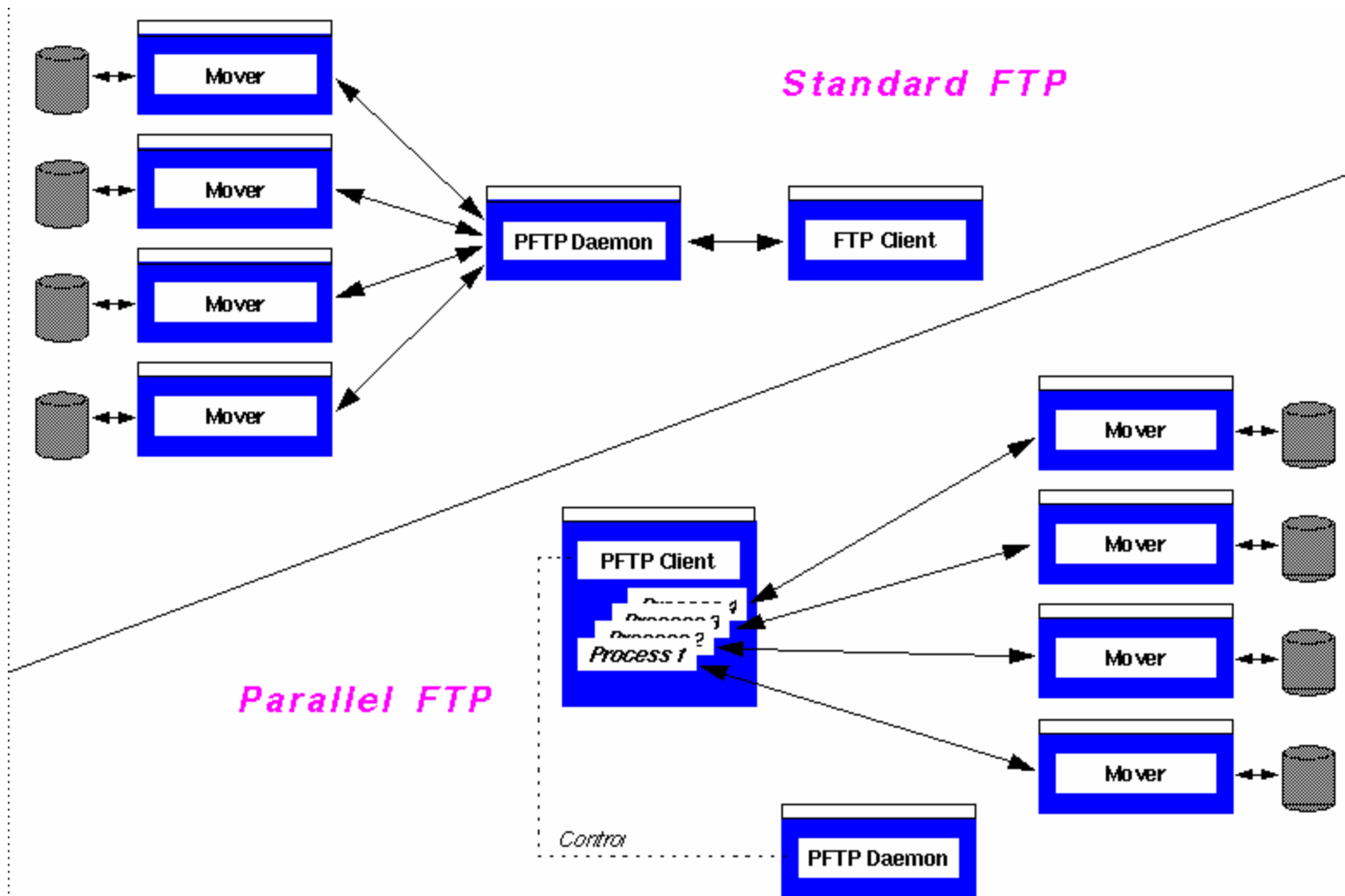
## Campus-wide network file servers

Indiana University distributed data management architecture



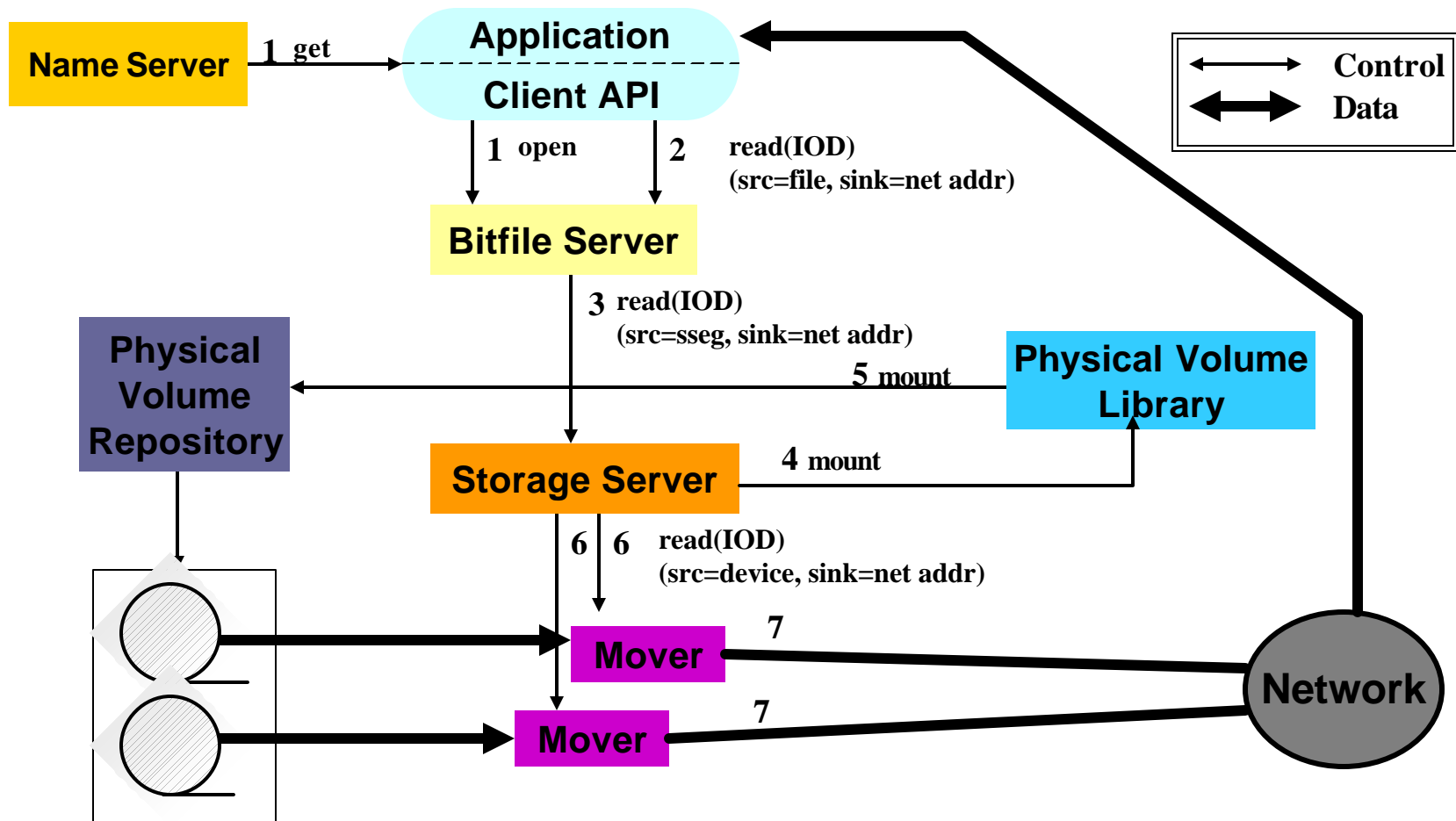
- HPSS serves as the deep, scalable store for the distributed multi-campus file system
- Serves the entire Indiana University community (>100,000 users)
- Plans to extend the architecture to collaborating institutions
- Will link multiple HPSS locations with a common name server and name space

# Standard vs. Parallel FTP Data Paths



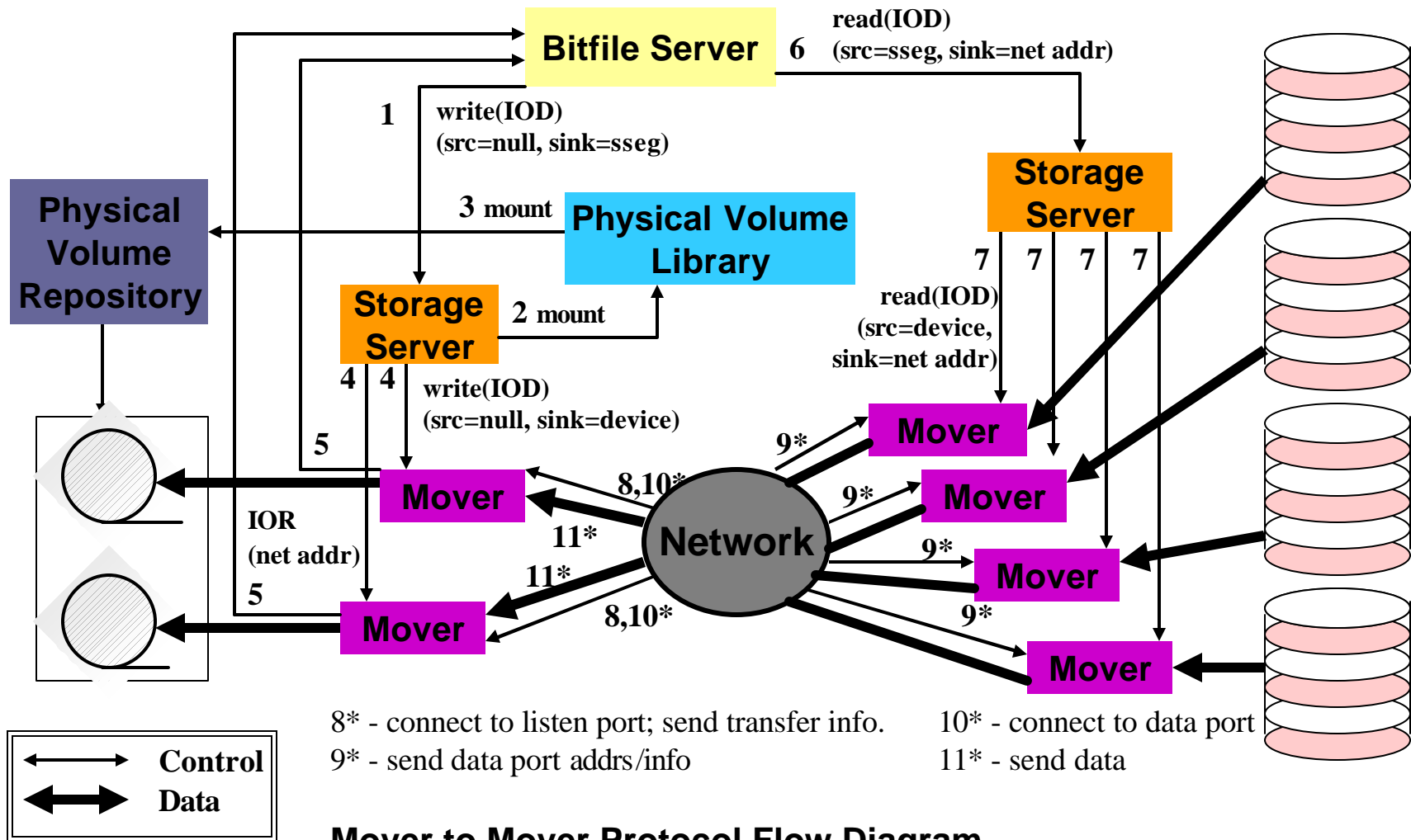


# Control and Data Flow



Simple Parallel Read Flow Diagram

# Control and Data Flow (cont'd)



# **HPSS Metrics**

- ♦ 420,000 executable lines of code in current release
- ♦ 200+ person years of development and test since inception - IBM, US Department of Energy
- ♦ 28 people currently working HPSS development and test - IBM, US Department of Energy
- ♦ 40+ Installations in US, Europe, Japan
- ♦ Quality
  - ♦ SEI Level 3 assessment
  - ♦ R&D 100 award (R&D Magazine, 1997)
  - ♦ Formal development and test procedures based on NASA and DoD standards

*HPSS is sold and serviced by  
IBM Global Services - Federal*

For more information:

[www.clearlake.ibm.com/hpss](http://www.clearlake.ibm.com/hpss)

Harry Hulen [hulen@us.ibm.com](mailto:hulen@us.ibm.com)